Appendix 1

Methods

Collection of exhaled breath samples

Participants exhaled directly into two Tedlar® bags. The volume of Bag 1 was 0.5 L (SKC Inc. Cat#232-01) and the volume of Bag 2 was 1 L (SKC Inc. Cat#232-02). Participants were instructed to take a deep inhalation and exhale ~150–300 mL of breath into the 0.5 L bag (Bag 1) to compensate for dead space. Immediately, the participant then inflated the 1 L bag (Bag 2) using the rest of the exhaled breath. All collected breath samples were delivered at room temperature to the Johns Hopkins Bloomberg School of Public Health (JHBSPH) research lab for VOC analysis within 2 h after collection (whenever possible). The balloon was measured within 24 h of breath by the lab in 166 (74%) patients. All but 3 balloons were read within 24 h of lab acquisition. Only data collected from Bag 2 were used for the analyses presented here.

VOC analysis with thermal desorption gas chromatography-mass spectrometry (TD-GC-MS)

VOCs in the exhaled breath were analyzed using TD and GC-MS. A multiple channel thermal desorption system (UNITY-xrTM) with an auto-sampler (CIA Advantage- xrTM both from Markes International, Inc., UK) was used to sample 100 mL of exhaled breath from each of the Tedlar bags at a flow rate of 50 mL/min and flow path temperature of 150 ℃. Helium was used as the carrier gas at a constant pressure of 5 Pounds per Square Inch (PSI); the sample was directly injected from the TD unit into the gas chromatograph for analysis.

Chromatographic analysis was performed using a Trace GC-Ultra gas chromatograph attached to an ISQ Mass Spectrometer (GC-MS, Thermo Scientific). VOC compounds were separated with a 30-m column × 0.25-mm internal diameter and 1.40 µm film thickness (Cat# 19915, Rtx-VMS, Restek Corp, U.S). The oven temperature was set on a gradient to achieve optimal separation of the analytes at an initial temperature of 35 ℃ with 1 min hold; the temperature rate was increased by 5 ℃/min to reach 100 ℃ followed by a final temperature ramp of 50 ℃/min to 240 ℃.

Calibration curves and quality control

Thirteen previously reported VOCs, representing different chemical groups, were selected for quantitative analysis; details are provided in *Table S1* in the supplementary materials. Clean and humidified air was injected into a subset (10%) of bags to evaluate measurement background. For each selected VOC, a five-point calibration curve was generated by spiking reagent-grade standards into Tedlar® bags in concentrations ranging from 0.390 µg/mL to 4,000 µg/mL using methanol as solvent. Exactly 1 µL aliquot of each standard was injected into five different bags filled with 1 L of pure Nitrogen, diluting the concentration of the analyte by 1,000×. Five calibration curves for each VOC were generated, and their average slope and intercept were used to quantify concentrations from participant samples.

Ten blanks were prepared by inflating Tedlar bags with clean and humidified air to evaluate measurement background. The lowest standard of each VOC was prepared at least five times and injected into the GC-MS to calculate the limit of detection. The limit of detection (LOD) for each chemical was calculated by multiplying the standard deviation of those low analytical standard replicates by 3 (LOD = StDev × 3). All lab analysts were blinded to study participants' status and information. Standardized procedures were used for performing and documenting lab operations, including sample management (login, registration integrity, life cycle tracking), chain of custody, inventory and storage management.

Statistical analysis

The main goals of the analysis were to: (I) detect whether VOC are predictive of S1LC; (II) quantify predictive performance of individual compounds; and (III) identify the best subset of VOC predictors for S1LC cases. Since many VOCs were below the limit of detection (LOD) for a large percentage of observations, only four VOCs with less than 10% data below the LOD were used in analyses that used concentrations as continuous variables. Each concentration was log-transformed to reduce the skewness of the marginal distributions. Additional models were fit with each individual VOC being above/below the LOD as a predictor and S1LC as an outcome using univariate logistic regression analysis. According to the study analytic protocol described before data were collected, the first 30 groups of matched cases and controls, determined by case enrollment time, were used for training and the last 58 groups were used for testing. Analyses were conducted by combining the two types of

https://dx.doi.org/10.21037/tlcr-21-953

controls, whenever they were both available. The reason for conducting matching during the design and implementation phase of the study was to improve covariate balance. However, all analyses were marginal and did not incorporate the matched pair design to ensure generalizability and practicality of findings. Indeed, in practice matched controls would typically not be available when an individual takes a screening test.

Each model was fit to the training data, and then applied to: (I) the testing; and (II) the combined testing and training data. All analyses were performed in the R statistical software (21). To detect statistically significant differences between VOC breath concentrations in S1LC and controls, two sample unpaired *t*-tests were performed using the R function *t*-test(). This was conducted only on the four VOCs with small percentage of concentrations below the LOD. Classification tests using thresholds of the statistically significant VOC were developed based on the 10^{th}, 25^{th} and 50^{th} percentiles of VOC concentrations in the training data of controls. Univariate and multivariate forward selection logistic regression models were fit using the glm() function in R. Forward selection was used to identify the combination of most predictive VOCs. Selection of VOCs were based on the improvement in the receiver operating characteristic area under the curve (AUC) in the training data. Each VOC was added to the model, the AUC was estimated again on the training set, and the VOC with the highest AUC was incorporated into the model. For each selected model the AUC on the test data was computed. Missing observations were excluded in each candidate model when individual VOCs were below the LOD.

Results

Figure S1 displays the Acetoin concentration for each S1LC case (red dots) and control (green dots) in the test data. On each vertical line there are either: (I) two dots (one red and one green), when the group contains a biopsy-confirmed S1LC case and a type 1 control; or (II) three dots (one red and two green) when the group contains a biopsy-confirmed S1LC case, a type 1 control, and a type 2 control. The dashed horizontal lines correspond to the classification thresholds based on the distribution of Acetoin concentration in controls in the training data set: 10th percentile shown in black (0.026 mg/L), 25th percentile shown in blue (0.044 mg/L) and 50th percentile (0.098 mg/L) shown in magenta. For each threshold, study participants with concentrations below the corresponding line are classified as cases and above the line as controls. The color of the dots is the true S1LC case status (red S1LC case, green control), while the position of the dot relative to one of the horizontal lines is the prediction of S1LC case status (below cancer, above control). This Figure provides the visual tradeoff in terms of false positives and false negative predictions as a function of the threshold on Acetoin concentrations.

https://dx.doi.org/10.21037/tlcr-21-953

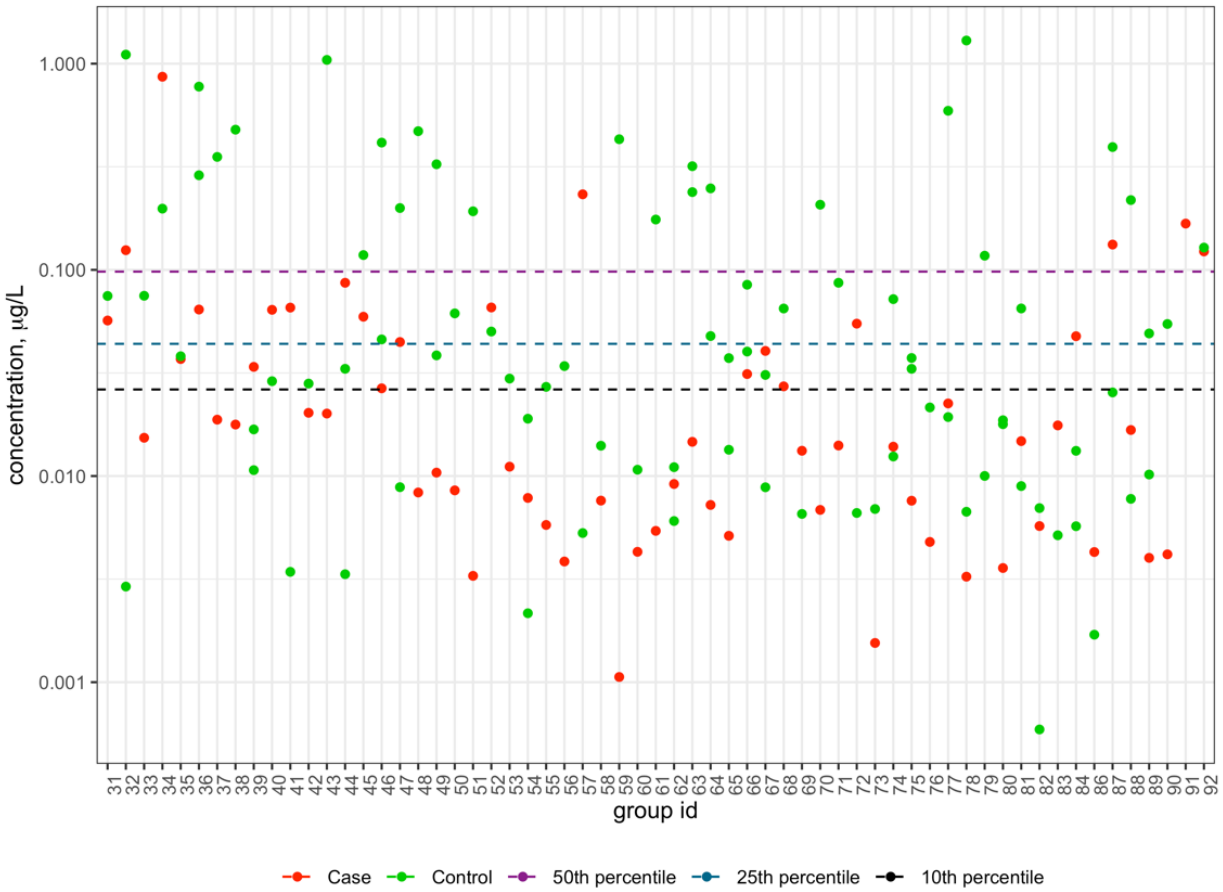

Figure S1 Classification based on Acetoin concentration threshold using the test data. The x-axis is the group number (starting at 31 because the first 30 groups are for training), each group with either two or three study participants. The y-axis is labeled on the concentration scale (mg/L), but data are $\log_{10}$ transformed in the model. Each point is a study participant (red S1LC case, green control). Horizontal lines correspond to three thresholds based on percentiles of the Acetoin concentrations distribution in all training data controls: 10th (0.026 mg/L, shown in black), 25th (0.044 mg/L, shown in blue) and 50th (0.098 mg/L, shown in magenta). For each threshold, participants below the line are classified as cases and above the line as controls.

https://dx.doi.org/10.21037/tlcr-21-953

Table S1 Chemicals selected for quantification, and references

No.	CAS	Names	Classification	Studies reporting the chemical
1	78-93-3	2-butanone (MEK)	Ketone	(29-34)
2	513-86-0	Acetoin	Ketone	(33,34)
3	108-88-3	Toluene	Aromatic hydrocarbon	(35,36)
4	107-87-9	2-Pentanone	Ketone	(31,32,37,38)
5	562-49-2	3,3-dimethyl-pentane	Alkane	(36)
6	123-51-3	3-methyl butanol	Alcohol	(39)
7	626-93-7	2-Hexanol	Alcohol	(39)
8	112-40-3	Dodecane	Alkane	(39)
9	66-25-1	Hexanal	Aldehyde	(39-42)
10	108-94-1	Cyclohexanone	Ketone	(39)
11	111-71-7	Heptanal	Aldehyde	(39,40,42)
12	100-41-4	Ethylbenzene	Aromatic hydrocarbon	(32,39,42,43)
13	99-87-6	p-Cymene	Aromatic hydrocarbon	(36)

References

29. Gordon SM, Szidon JP, Krotoszynski BK, et al. Volatile organic compounds in exhaled air from patients with lung cancer. Clin Chem 1985;31:1278-82.
30. Bajtarevic A, Ager C, Pienz M, et al. Noninvasive detection of lung cancer by analysis of exhaled breath. BMC Cancer 2009;9:348.
31. Ligor M, Ligor T, Bajtarevic A, et al. Determination of volatile organic compounds in exhaled breath of patients with lung cancer using solid phase microextraction and gas chromatography mass spectrometry. Clin Chem Lab Med 2009;47:550-60.
32. Buszewski B, Ligor T, Jezierski T, et al. Identification of volatile lung cancer markers by gas chromatography-mass spectrometry: comparison with discrimination by canines. Anal Bioanal Chem 2012;404:141-6.
33. Fu XA, Li M, Knipp RJ, et al. Noninvasive detection of lung cancer using exhaled breath. Cancer Med 2014;3:174-81.
34. Schumer EM, Trivedi JR, van Berkel V, et al. High sensitivity for lung cancer detection using analysis of exhaled carbonyl compounds. J Thorac Cardiovasc Surg 2015;150:1517-22; discussion 1522-4.
35. Poli D, Carbognani P, Corradi M, et al. Exhaled volatile organic compounds in patients with non-small cell lung cancer: cross (2sectional and nested short-term follow-up study. Respir Res 2005;6:71.
36. Peng G, Hakim M, Broza YY, et al. Detection of lung, breast, colorectal, and prostate cancers from exhaled breath using a single array of nanosensors. Br J Cancer 2010;103:542-51.
37. Ulanowska A, Kowalkowski T, Trawińska E, et al. The application of statistical methods using VOCs to identify patients with lung cancer. J Breath Res 2011;5:046008.
38. Li M, Yang D, Brock G, et al. Breath carbonyl compounds as biomarkers of lung cancer. Lung Cancer 2015;90:92-7.
39. Handa H, Usuba A, Maddula S, et al. Exhaled breath analysis for lung cancer detection using ion mobility spectrometry. PLoS One 2014;9:e114555.
40. Chen X, Xu F, Wang Y, et al. A study of the volatile organic compounds exhaled by lung cancer cells in vitro for breath diagnosis. Cancer 2007;110:835-44.
41. Fuchs P, Loeseken C, Schubert JK, et al. Breath gas aldehydes as biomarkers of lung cancer. Int J Cancer 2010;126:2663-70.
42. Poli D, Goldoni M, Corradi M, et al. Determination of aldehydes in exhaled breath of patients with lung cancer by means of on-fiber-derivatisation SPME-GC/MS. J Chromatogr B Analyt Technol Biomed Life Sci 2010;878:2643-51.
43. Rudnicka J, Kowalkowski T, Ligor T, et al. Determination of volatile organic compounds as biomarkers of lung cancer by SPME-GC-TOF/MS and chemometrics. J Chromatogr B Analyt Technol Biomed Life Sci 2011;879:3360-6.

https://dx.doi.org/10.21037/tlcr-21-953

Table S2 The number of participants with VOCs below limit of detection in the combined data set and stratified by testing and training groups

Compound	Limit of detection (μg/L)	Combined data (N=225)	Training data		Test data	
			Case (N=30)	Control (N=51)	Case (N=58)	Control (N=86)
3,3-dimethyl pentane	0.00181	219 (97.3%)	26 (86%)	38 (74%)	56 (96%)	80 (94%)
2-Butanone	0.00815	218 (96.9%)	28 (94%)	48 (94%)	57 (98%)	85 (98%)
2-Pentanone	0.00130	17 (7.6%)	2 (6%)	3 (6%)	3 (6%)	9 (10%)
Toluene	0.01854	222 (98.7%)	28 (94%)	50 (98%)	58 (100%)	86 (100%)
3-Methyl-1-Butanol	0.00181	200 (88.9%)	26 (86%)	38 (74%)	56 (96%)	80 (94%)
Acetoin	0.00037	7 (3.1%)	2 (6%)	2 (4%)	1 (2%)	2 (2%)
2-Hexanol	0.00199	172 (76.4%)	18 (60%)	25 (50%)	51 (88%)	78 (90%)
Hexanal	0.00334	209 (92.9%)	27 (90%)	43 (84%)	56 (96%)	83 (96%)
Ethylbenzene	0.00074	217 (96.4%)	28 (94%)	48 (94%)	56 (96%)	85 (98%)
Heptanal	0.00023	19 (8.4%)	7 (24%)	6 (12%)	2 (4%)	4 (4%)
Cyclohexanone	0.00581	148 (65.8%)	6 (20%)	15 (30%)	52 (90%)	75 (88%)
p-Cymene	0.00011	126 (56.0%)	19 (64%)	19 (38%)	41 (70%)	47 (54%)
Dodecane	0.00002	4 (1.8%)	1 (4%)	0 (0%)	2 (4%)	1 (2%)

Table S3 Correlations of log concentrations of quantifiable VOCs that were above the limit of detection in more than 10% of the samples

	2-Pentanone	Acetoin	Dodecane	Heptanal
2-Pentanone	1.000	0.110	0.656	0.446
Acetoin	0.110	1.000	0.374	0.369
Dodecane	0.656	0.374	1.000	0.488
Heptanal	0.446	0.369	0.488	1.000

https://dx.doi.org/10.21037/tlcr-21-953